



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/909,543	07/19/2001	Sheng Li	3442P014	1957
8791	7590	02/08/2006	EXAMINER	
BLAKELY SOKOLOFF TAYLOR & ZAFMAN 12400 WILSHIRE BOULEVARD SEVENTH FLOOR LOS ANGELES, CA 90025-1030			SINGH, RACHNA	
			ART UNIT	PAPER NUMBER
			2176	

DATE MAILED: 02/08/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	09/909,543	LI, SHENG	
	<b>Examiner</b>	<b>Art Unit</b>	
	Rachna Singh	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

1) Responsive to communication(s) filed on 17 November 2005.

2a) This action is **FINAL**.                    2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

4) Claim(s) 1-18 and 24-28 is/are pending in the application.

4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.

5) Claim(s) \_\_\_\_\_ is/are allowed.

6) Claim(s) 1-18 and 24-28 is/are rejected.

7) Claim(s) \_\_\_\_\_ is/are objected to.

8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on \_\_\_\_\_ is/are: a) accepted or b) objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a) All    b) Some \* c) None of:

1. Certified copies of the priority documents have been received.
2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)	4) <input type="checkbox"/> Interview Summary (PTO-413)
2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)	Paper No(s)/Mail Date. _____ .
3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____ .	5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)
	6) <input type="checkbox"/> Other: _____ .

### **DETAILED ACTION**

1. This action is responsive to communications: RCE filed 11/17/05. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 11/17/05 has been entered.
2. Claims 1-18 and 24-28 are pending. Claims 19-23 have been cancelled. Claims 1, 7, 10, 12, and 14 are independent claims.

#### ***Claim Rejections - 35 USC § 112***

3. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.
4. Claims 1, 7, 10, 12, and 14 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. Claims recite a limitation "dividing the audio sequence into a plurality of equally-sized audio data groups". The specification does not appear to describe the claimed subject matter. Applicant is requested to indicate portions of the specification where dividing the audio sequence into equally sized audio data groups is disclosed.

***Claim Rejections - 35 USC § 103***

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 1-6, 10, 12, 14, and 24-28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Witteman, US 2002/0055950 A1, 5/9/02 (filed 4/23/01, continuation of an application filed 12/23/98).

In reference to claims 1, 10, 12, and 14, Witteman teaches synchronizing audio and text of multimedia segments. See abstract. Compare to "***A method for synchronizing multimedia data having at least audio and text sequences***".

Witteman teaches the following:

-Separating the audio component and closed caption component from a single stream. Generating an audio pattern representative of the start of the multimedia segment; locating the audio pattern in the audio component; generating a concluding audio pattern representative of the end of the multimedia segment; locating the concluding audio pattern in the audio component; identifying the multimedia segment between the audio patterns. See page 1, paragraphs [0005]-[0009]. Determining the start of the audio block, indexing the audio block, and sending the audio block to an information store. See page 2, paragraphs [0027]-[0029], page 3, paragraph [0032], and figure 3.

Witteman discloses temporally aligning the text with the audio pattern in the audio

component. See page 1, paragraph [0010] and figure 3, elements 444, 446, and 448 which illustrate temporally aligning (in seconds) the audio information with the text information using text marks (in seconds). Compare to "***dividing the audio sequence into a plurality of equally-sized audio data groups; matching each audio data group of said plurality of audio data groups to a nearest time mark within a discrete series of time marks separated by a predefined time period***". Temporally aligning the audio and text information in seconds is "equally" dividing the groups into equally sized segments.

-Comparing the text against one or more keywords delimiting the multimedia segment and temporally aligning the text with the audio pattern in the audio component. See pages 1-2 and figure 3. Compare to "***associating each audio data group. . .in the text sequence***".

Wittman teaches associating the audio pattern to words in a text sequence using a temporal alignment; however, he does not state that a number is used to associate the word to the audio group and each number is uniquely identifying a particular word. The "number of the word" is used to put the words of a text sequence in order. Wittman teaches that the text in the closed caption components are aligned temporally. See figure 3, 448 illustrating time in seconds associated with the various audio and closed-caption (i.e. text) information. Applicant's specification on pages 5-7 recites, "*the words in the text sequence may then be synchronized to the audio data groups by linking the word number with each audio data group. A special word number may be used to indicate that the text should not be advanced when the word audio*

*portion is longer than the audio data group size or when the current audio data group has a sound gap . . . the word ordinal number 302 represents the order of a word within a text sequence.”* It would have been obvious to a person of ordinary skill in the art at the time of the invention to equate Wittelman’s temporal alignment to the “numbering” the words of a text sequence since both the temporal alignment and the numbering of the words allow the text or phrase to be ordered in a sequential manner which then allows each word of text sequence to be associated with a specific audio group. As further illustrated in figure 3, Wittelman teaches associating the audio pattern to words in a text sequence using a temporal alignment where the temporal number (448) are used to illustrate time in seconds associated with the various audio and closed-caption (i.e. text) information.

In reference to claims 2, 3, and 6, Wittelman teaches generating an audio pattern representative of the start of the multimedia segment; locating the audio pattern in the audio component; generating a concluding audio pattern representative of the end of the multimedia segment; locating the concluding audio pattern in the audio component; identifying the multimedia segment between the audio patterns. See page 1, paragraphs [0005]-[0009]. Determining the start of the audio block, indexing the audio block, and sending the audio block to an information store. See page 2, paragraphs [0027]-[0029], page 3, paragraph [0032], and figure 3. The start and end of the multimedia segment determine the size of the audio frame. The audio pattern is segmented accordingly. The size of the audio segment is not limited in any manner and could include a size of 100 milliseconds. See figure 3. Wittelman discloses temporally

aligning the text with the audio pattern in the audio component. See page 1, paragraph [0010] and figure 3, elements 444, 446, and 448 which illustrate temporally aligning the audio information with the text information using text marks (in seconds).

In reference to claims 4 and 5, Witteman's system temporally aligns the text to the audio pattern. If there is no text for the selected audio component, then the audio component is temporally assigned to nothing except the time. See figure 3.

In reference to claims 24-28, Witteman discloses temporally aligning the text with the audio pattern in the audio component. See page 1, paragraph [0010] and figure 3, elements 444, 446, and 448 which illustrate temporally aligning the audio information with the text information using text marks (in seconds).

7. Claims 7-9, 11, 13, and 15-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Witteman, US 2002/0055950 A1, 5/9/02 (filed 4/23/01, continuation of an application filed 12/23/98) in view of Ishii, US 6,778,493 B1, 8/17/04 (filed 2/7/00).

In reference to claims 7-9, Witteman teaches synchronizing audio and text of multimedia segments. See abstract. Compare to "***A method for synchronizing a text sequence with an audio sequence***". Witteman teaches the following:

-Separating the audio component and closed caption component from a single stream. Generating an audio pattern representative of the start of the multimedia segment; locating the audio pattern in the audio component; generating a concluding audio pattern representative of the end of the multimedia segment; locating the concluding audio pattern in the audio component; identifying the multimedia segment between the audio patterns. See page 1, paragraphs [0005]-[0009]. Determining the start of the

audio block, indexing the audio block, and sending the audio block to an information store. See page 2, paragraphs [0027]-[0029], page 3, paragraph [0032], and figure 3. Compare to "**arranging the audio sequence into a plurality of audio data groups; synchronizing a current audio data group of said at least one audio data group to a nearest time mark**". Temporally aligning the audio and text information in seconds is "equally" dividing the groups into equally sized segments.

-Comparing the text against one or more keywords delimiting the multimedia segment and temporally aligning the text with the audio pattern in the audio component. See pages 1-2 and figure 3.

Witteman teaches associating the audio pattern to words in a text sequence using a temporal alignment; however, he does not state that a number is used to associate the word to the audio group and each number is uniquely identifying a particular word. The "number of the word" is used to put the words of a text sequence in order. Witteman teaches that the text in the closed caption components are aligned temporally. See figure 3, 448 illustrating time in seconds associated with the various audio and closed-caption (i.e. text) information. Applicant's specification on pages 5-7 recites, "*the words in the text sequence may then be synchronized to the audio data groups by linking the word number with each audio data group. A special word number may be used to indicate that the text should not be advanced when the word audio portion is longer than the audio data group size or when the current audio data group has a sound gap . . . the word ordinal number 302 represents the order of a word within a text sequence.*" It would have been obvious to a person of ordinary skill in the art at

the time of the invention to equate Witteman's temporal alignment to the "numbering" the words of a text sequence since both the temporal alignment and the numbering of the words allow the text or phrase to be ordered in a sequential manner which then allows each word of text sequence to be associated with a specific audio group. As further illustrated in figure 3, Witteman teaches associating the audio pattern to words in a text sequence using a temporal alignment where the temporal number (448) are used to illustrate time in seconds associated with the various audio and closed-caption (i.e. text) information.

Most modern Wide Area Network (WAN) protocols at the time of the invention were based on packet-switching technologies. See figure 5. Witteman does not explicitly teach the packetization of the audio groups and words; however, Ishii illustrates this feature. Ishii teaches real-time media content synchronization and transmission in packet network apparatus and method. Ishii's system teaches transmitting and synchronizing multimedia content for generating a multimedia packet having multimedia audio/visual information and for transmitting the multimedia packet. See abstract and column 3-4. It would have been obvious to a person of ordinary skill in the art at the time of the invention to incorporate the packetization of audio and text for delivery over a network since it was well known in the art at the time of the invention to synchronize and transmit multimedia data streams from one or more sources over a packet-based system to multiple receivers since it would allow multimedia contents to be played in a synchronized manner. See pages 1-4 of Ishii.

In reference to claim 11, most modern Wide Area Network (WAN) protocols were based on packet-switching technologies. See figure 5. Witteman's system could include the packetization of the audio groups and words. Ishii further illustrates this feature. Ishii teaches real-time media content synchronization and transmission in packet network apparatus and method. Ishii's system teaches transmitting and synchronizing multimedia content for generating a multimedia packet having multimedia audio/visual information and for transmitting the multimedia packet. See abstract and column 3-4. It would have been obvious to a person of ordinary skill in the art at the time of the invention to incorporate the packetization of audio and text for delivery over a network since it was well known in the art at the time of the invention to synchronize and transmit multimedia data streams from one or more sources over a packet-based system to multiple receivers since it would allow multimedia contents to be played in a synchronized manner. See pages 1-4 of Ishii.

In reference to claim 13, most modern Wide Area Network (WAN) protocols were based on packet-switching technologies. See figure 5. Witteman's system could include the packetization of the audio groups and words. Ishii further illustrates this feature. Ishii teaches real-time media content synchronization and transmission in packet network apparatus and method. Ishii's system teaches transmitting and synchronizing multimedia content for generating a multimedia packet having multimedia audio/visual information and for transmitting the multimedia packet. See abstract and column 3-4. It would have been obvious to a person of ordinary skill in the art at the time of the invention to incorporate the packetization of audio and text for delivery over

a network since it was well known in the art at the time of the invention to synchronize and transmit multimedia data streams from one or more sources over a packet-based system to multiple receivers since it would allow multimedia contents to be played in a synchronized manner. See pages 1-4 of Ishii.

In reference to claims 15-18, Witteman teaches comparing the text against one or more keywords delimiting the multimedia segment and temporally aligning the text with the audio pattern in the audio component. See pages 1-2 and figure 3. Most modern Wide Area Network (WAN) protocols were based on packet-switching technologies. See figure 5. Thus Witteman's system inherently includes packetizing of the audio groups and words/text sequences. Furthermore, Witteman discloses a computer system with a file sharing protocol on top of its TCP/IP protocol (most TCP/IP were based on packet-switching technologies at the time of the invention). See page 5. Most modern Wide Area Network (WAN) protocols were based on packet-switching technologies. See figure 5. Ishii further illustrates this feature. Ishii teaches real-time media content synchronization and transmission in packet network apparatus and method. Ishii's system teaches transmitting and synchronizing multimedia content for generating a multimedia packet having multimedia audio/visual information and for transmitting the multimedia packet. See abstract and column 3-4. It would have been obvious to a person of ordinary skill in the art at the time of the invention to incorporate the packetization of audio and text for delivery over a network since it was well known in the art at the time of the invention to synchronize and transmit multimedia data streams from one or more sources over a packet-based system to multiple receivers since it

would allow multimedia contents to be played in a synchronized manner. See pages 1-4 of Ishii.

***Response to Arguments***

8. Applicant's amendments filed 11/17/05 have been reconsidered, but are not persuasive.

Applicant argues that Witteman does not discloses "assigning a number to each of a plurality of words in a text sequence, each number uniquely identifying a particular word", "synchronizing an audio data group to a nearest time mark within a series of time marks spaced according to a predefined temporal arrangement", or "associating an audio data group to a number of a word in a text sequence corresponding to audio content contained within the audio data group". Applicant further argues that Witteman does not teach a temporal arrangement for synchronizing audio data groups. Examiner respectfully disagrees. Witteman explicitly states temporally aligning the text with the audio pattern in the audio component. See page 1, paragraph [0010] and figure 3, elements 444, 446, and 448 which illustrate temporally aligning the audio information with the text information using text marks (in seconds). Furthermore, Witteman discloses associating the audio data group to words in the text sequence. See pages 1-2 and figure 3. Witteman teaches associating the audio pattern to words in a text sequence using a temporal alignment; however, he does not state that a number is used to associate the word to the audio group and each number is uniquely identifying a particular word. The "number of the word" is used to put the words of a text sequence in order. Witteman teaches that the text in the closed caption components are aligned

temporally. See figure 3, 448 illustrating time in seconds associated with the various audio and closed-caption (i.e. text) information. Applicant's specification on pages 5-7 recites, "*the words in the text sequence may then be synchronized to the audio data groups by linking the word number with each audio data group. A special word number may be used to indicate that the text should not be advanced when the word audio portion is longer than the audio data group size or when the current audio data group has a sound gap . . . the word ordinal number 302 represents the order of a word within a text sequence.*" It would have been obvious to a person of ordinary skill in the art at the time of the invention to equate Witteman's temporal alignment to the "numbering" the words of a text sequence since both the temporal alignment and the numbering of the words allow the text or phrase to be ordered in a sequential manner which then allows each word of text sequence to be associated with a specific audio group. As further illustrated in figure 3, Witteman teaches associating the audio pattern to words in a text sequence using a temporal alignment where the temporal number (448) are used to illustrate time in seconds associated with the various audio and closed-caption (i.e. text) information.

Applicant argues Witteman does not teach dividing audio groups into equally sized groups; however, Witteman teaches temporally aligning the audio and text information in seconds. Dividing according to seconds (as illustrated in figure 3) is "equally" dividing the sequence into equally sized segments.

In view of the comments above, the rejection is maintained.

### ***Conclusion***

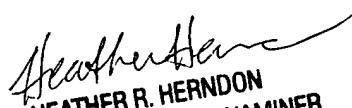
9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Hu, Michael and Ye Jian. "Multimedia Description Framework (MDF) for Content Description of Audio/Video Documents", ACM 1999.

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Rachna Singh whose telephone number is 571-272-4099. The examiner can normally be reached on M-F (8:30AM-6:00PM). If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 571-272-4090.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

RS  
01/26/06

  
HEATHER R. HERNDON  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100